

VŠB – Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky

**VYUŽITÍ MODERNÍCH STATISTICKÝCH METOD
PRO ANALÝZU NEŽÁDOUCÍCH ÚČINKŮ SPOJENÝCH
S RADIOTERAPIÍ KARCINOMU PROSTATY**

Autoreferát dizertační práce

ING. MARTINA LITSCHMANNOVÁ

Studijní program: P2649 Elektrotechnika
Studijní obor: 2612V045 Technická kybernetika
Školitel: prof. Ing. Radim Briš, CSc.

Ostrava 2011

ABSTRAKT

Dizertační práce na téma Využití moderních statistických metod pro analýzu nežádoucích účinků spojených s radioterapií karcinomu prostaty ukazuje, a v konkrétním případě řeší, nedostatky, k nimž dochází při sběru medicínských dat, a dále se zaměřuje na způsoby odhadu nejistot při určování incidence sledovaných jevů a na korelační analýzu zaměřenou na specifická biomedicínská data.

Vlastní práce obsahuje 8 kapitol. Úvodní část (kapitoly 2-4) je věnována základním teoretickým poznatkům o karcinomu prostaty a jeho léčbě. Z těchto poznatků je pak vycházeno při návrhu software, který slouží k vytváření a shromažďování lékařských záznamů o pacientech, o průběhu jejich hormonální léčby a průběhu a následcích radioterapie (kapitola 5). Kapitola 6 představuje a srovnává různé typy intervalových odhadů parametru binomického rozdělení s ohledem na využití pro analýzu neurčitosti v biomedicínských datech. Volně šiřitelný výpočetní applet umožňující výpočet analyzovaných intervalových odhadů je součástí příloženého CD. Kapitola 7 se věnuje korelační analýze, zejména pak biseriálním, polyseriálním a polychorickým korelacím. Součástí příloženého CD jsou výpočetní applety pro snadný výpočet specifických korelačních koeficientů mezi dvojicemi spojitých proměnných, dvojicemi ordinálních proměnných a mezi spojitou proměnnou a ordinální proměnnou vzniklou diskretizací skryté spojitě proměnné. Kapitola 8 je věnována aplikaci problematiky řešené v kapitolách 6 a 7 na konkrétní onkologická data.

ABSTRACT

PhD thesis on use of modern statistical methods for analysis of side effects associated with radiotherapy for prostate cancer shows, and in the particular case solves, deficiencies that arise during the collection of medical data, and focuses on ways to estimate the uncertainties in determining the incidence of observed events and the correlation analysis focused on specific biomedical data.

Dissertation consists of 8 chapters. The introductory part (Chapters 2-4) is devoted to the basic theoretical knowledge about prostate cancer and its treatment, which is then going out to design software for creating and collecting medical records of patients on the course of hormone treatment and radiotherapy course and consequences (Chapter 5). Chapter 6 presents and compares different types of confidence interval for parameter of binomial distribution with respect to the analysis of biomedical data. Computing freeware applet allowing calculation of the analyzed confidence intervals is included in the attached CD. Chapter 7 is devoted to correlation analysis, especially biserial, polyserial and polychoric correlations. Parts of the attached CD are computer applets for easy calculation of the correlation between two continuous variables, two ordinal variables and between continuous variable and ordinal variable resulting discretization hidden continuous variable. Chapter 8 is devoted to application issues dealt with in Chapters 6 and 7 on the specific cancer data.

OBSAH

1	ÚVOD	5
1.1	CÍLE DIZERTAČNÍ PRÁCE	5
1.2	VZTAH DIZERTAČNÍ PRÁCE K OBORU KYBERNETIKA	5
2	KARCINOM PROSTATY A JEHO LÉČBA	6
3	ORGANIZACE SBĚRU DAT.....	7
3.1	SOFTWARE MEDICALSUITE.....	7
4	INTERVALOVÉ ODHADY INCIDENCE JEDNOTLIVÝCH TYPŮ TOXICITY	9
4.1	POROVNÁNÍ ANALYZOVANÝCH INTERVALOVÝCH ODHADŮ	10
4.1.1	Porovnání analyzovaných liberálních intervalových odhadů.....	11
4.1.2	Porovnání analyzovaných konzervativních intervalových odhadů	14
4.2	SOFTWARE PRO VÝPOČET INTERVALOVÉHO ODHADU PARAMETRU BINOMICKÉHO ROZDĚLENÍ	15
5	IDENTIFIKACE VLIVU RADIOTERAPIE A DALŠÍCH FAKTORŮ NA INTENZITU NEŽ. ÚČINKŮ	
RADIOTERAPIE		16
5.1	SOFTWARE PRO VÝPOČET KORELAČNÍCH KOEFICIENTŮ	16
6	VYHODNOCENÍ INCIDENCE A PREDIKTORŮ TOXICITY PROVÁZEJÍCÍ RADIOTERAPII KARCINOMU	
PROSTATY		17
7	ZÁVĚRY PLYNOUCÍ Z PŘEDLOŽENÉ DIZERTAČNÍ PRÁCE	20
7.1	ZÁVĚRY PLYNOUCÍ PRO PRAXI	20
7.2	ZÁVĚRY PLYNOUCÍ PRO ROZVOJ VĚDY	22
7.3	MOŽNOSTI DALŠÍHO POKRAČOVÁNÍ V ŘEŠENÍ PROBLEMATIKY	23
	SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....	24
	LITERATURA	25
	CURICULUM VITAE	26
	VZDĚLÁNÍ	26
	PRAXE	26
	ODBORNÁ ČINNOST DOKTORANDKY	26
	ODBORNÁ ČINNOST S PŘÍMÝM VZTAHEM K DIZERTAČNÍ PRÁCI	26
	<i>Odborné publikace na Web Of Science</i>	<i>26</i>
	<i>Další odborné Publikace.....</i>	<i>26</i>
	<i>Přednášky.....</i>	<i>27</i>
	<i>Registrovaný software</i>	<i>27</i>
	<i>Účast na grantech.....</i>	<i>27</i>
	OSTATNÍ ODBORNÁ ČINNOST	27
	<i>Odborné publikace na Web Of Science</i>	<i>27</i>
	<i>Účast na grantech.....</i>	<i>27</i>
	<i>Odborné publikace.....</i>	<i>27</i>
	<i>Skripta 28</i>	
	<i>Přednášky.....</i>	<i>28</i>
	<i>Prezentace na konferencích</i>	<i>28</i>

1 ÚVOD

Na Katedře aplikované matematiky, FEI, VŠB-TU Ostrava se, zejména v souvislosti se spoluprací s Fakultní nemocnicí v Ostravě (FNO), setkáváme s požadavky na analýzu medicínských dat. Jde o oblast, která se určitým způsobem dotýká každého člověka. Lékaři řeší spoustu zajímavých, často i složitých problémů, které vedou na více či méně sofistikované statistické metody a modely. V letech 2006 – 2009 se ve FNO v souvislosti s léčbou karcinomu prostaty podrobilo 408 pacientů radioterapii. Ve spolupráci s MUDr. Ing. J. Cvekem, PhD. jsme se pokusili záznamy o těchto pacientech tvůrčím způsobem analyzovat a vyvodit relevantní závěry.

1.1 CÍLE DIZERTAČNÍ PRÁCE

Na základě proběhnuvší státní doktorské zkoušky byly stanoveny čtyři cíle předkládané dizertační práce s názvem Využití moderních statistických metod pro analýzu nežádoucích účinků spojených s radioterapií karcinomu prostaty:

1. Organizace sběru dat spojených s léčbou karcinomu prostaty a generování softwarové aplikace pro jejich předzpracování.
2. Analýza možných metod odhadu nejistot při určování incidence sledovaných jevů vycházející ze simulace pravděpodobnosti pokrytí konfidenčních intervalů. Generování výpočetního appletu pro kvantifikaci a vzájemné porovnání těchto nejistot.
3. Hlubková korelační analýza modifikovaná pro biomedicínská data a jejich specifické zvláštnosti. Generování výpočetního appletu umožňujícího vyčíslení specifických korelačních koeficientů.
4. Tvůrčí vyhodnocení záznamů o radioterapeutické léčbě karcinomu prostaty (2006 – 2009, MUDr. Ing. Jakub Cvek, PhD., Fakultní nemocnice Ostrava, Klinika onkologická), identifikace závislosti radioterapie a dalších faktorů v kontextu s jejími nežádoucími účinky.

1.2 VZTAH DIZERTAČNÍ PRÁCE K OBORU KYBERNETIKA

Kybernetika se zaměřuje především na to, jak systémy využívají informací o vazbě mezi svými vstupy a výstupy ke kontrole své vlastní funkce. Znalost systému je tvořena jejich modely (zjednodušenými výklady těchto systému) zanedbávajícími vlastnosti systému, které jsou pro účel, k němuž byl model vytvořen nepodstatné. Díky interdisciplinaritě, která je kybernetickému uvažování vlastní, je možné ji používat k pochopení a modelování libovolných systémů (technických, fyzikálních, biologických, ekologických, ...).

Pokud lékaři zvolí u pacienta určitý způsob léčby (vstup), pacient (systém) bude na základě svého stavu popsaného stavovými proměnnými (např. věk, krevní tlak, závažnost onemocnění, ...) reagovat určitým způsobem (výstup), např. objeví se u něj inkontinence. Ve většině případů mají lékaři jen malé povědomí o konkrétním mechanismu, který vede k pozorovanému efektu. Léčba spouští komplexní řetězec navzájem propojených reakcí zahrnujících různé orgány a části těla, nicméně jediné co můžeme pozorovat je konečný výsledek. Z tohoto hlediska považujeme pacienta za černou skříňku, tj. něco, co vstřebá vstupy a vyprodukuje výstupy, aniž bychom byli schopní vidět, co se děje mezi tím. Lékaře lze v tomto pohledu vnímat jako adaptivní regulátor nastavující své parametry na základě informace o reakci pacienta na léčbu.

Stavové a výstupní veličiny analyzovaného systému (pacientů) mají náhodnou povahu. Cíle předkládané dizertační práce směřují k:

- odhadu nejistot na výstupu systému (tj. odhadu nejistot při určování incidence míry nežádoucích účinků radioterapie),
- identifikaci vazeb mezi vstupy a výstupy (tj. identifikaci vlivu radioterapie na intenzitu nežádoucích účinků spjatých s touto léčbou) a identifikaci vazeb mezi stavovými proměnnými a výstupy (tj. identifikaci prediktorů toxicity urogenitálního a gastrointestinálního systému).

K identifikaci uvedených vazeb je využívána korelační analýza, která je považována za jednu ze základních metod užívaných pro stochastickou identifikaci systému, modifikována pro specifická biomedicínská data. Cíle 2 – 4 předkládané dizertační práce lze proto z pohledu kybernetiky považovat za stochastickou identifikaci systému.

2 KARCINOM PROSTATY A JEHO LÉČBA

Karcinom prostaty tvoří asi 4% všech nádorových onemocnění. V mužské populaci pak představuje tato rakovina 10% všech nádorů, podobně jako v jiných zemích západní Evropy. V České republice se jeho incidence pohybuje okolo 97 případů na 100 000 mužů, což je oproti počátku devadesátých let nárůst o 70%. Dobrou zprávou je, že tento nárůst lze připsat na vrub lepší diagnostice. Mortalita se dlouhodobě nemění a pohybuje se kolem 29 případů na 100 000 mužů.

Diagnostické metody vedoucí ke stanovení karcinomu prostaty zahrnují historicky nejstarší *vyšetření pohmatem přes konečník* (per rectum) a *stanovení sérové koncentrace PSA* [ng/ml] (prostate-specific antigen). **PSA** je produkován buňkami karcinomu prostaty a v současné době je považován za nejlepší sérový marker jak ke stanovení diagnózy karcinomu prostaty, tak ke zjištění stádia nemoci.

Na základě stádia nemoci (nízké, střední, vysoké riziko) pak lékaři rozhodují o další léčbě. Radioterapie nachází uplatnění v léčbě všech stádií tohoto onemocnění. Kurativní léčba je indikována

u lokalizovaného nebo lokálně pokročilého onemocnění (střední riziko). V těchto stádiích nabízí léčbu, která je, s výjimkou brachyterapie, neinvazivní, provázená nízkou akutní morbiditou a prakticky nulovou mortalitou, bez nutnosti celkové anestézie a hospitalizace. Paliativní léčba je používána tam, kde usilujeme o zmírnění obtíží vyvolaných nádorem či metastázou (vysoké riziko), ale nově též k ozaření existujících metastatických ložisek, která jsou ještě asymptomatická (nízké riziko). Od druhé základní léčebné modality, tj. chirurgické léčby, se liší nikoliv ve výsledcích, ale ve spektru nežádoucích účinků [6].

S vysokou dávkou záření je spojena i toxicita zdravých tkání. V případě radioterapie karcinomu prostaty jde o toxicitu oblasti rekta (gastrointestinálního traktu, GIT) a toxicitu oblasti močového měchýře (urogenitálního systému, UGS). **Akutně**, tj. během ozařování, se mohou objevit potíže, které po aplikaci terapie, zpravidla do 2 - 4 týdnů po ukončení ozařování, ustupují. Po ukončení ozařování (od 2-4 měsíců až do 2-3 let) se mohou objevit **pozdní** nežádoucí účinky léčby radioterapií. Tyto změny jsou trvalého charakteru a mohou významně ovlivnit život pacienta. Míra jak akutních, tak pozdních účinků radioterapie na námi sledované orgány se hodnotí na 5-ti bodové škále podle mezinárodně uznávaného skórovacího systému.

3 ORGANIZACE SBĚRU DAT

Vyšetřování pacienti představují pro lékaře objemný zdroj dat, jejichž analýza vyžaduje sofistikované přístupy. Při samotném sběru dat je třeba mít na mysli, že data se jednou budou statisticky zpracovávat. I v této fázi návrhu experimentu je vhodná spoluúčast datového analytika. V případě, že lékaři nemají k dispozici profesionální software pro zápis sledovaných proměnných pro jednotlivé subjekty, používají nejčastěji MS Excel. Za podstatné nedostatky komplikující následnou analýzu lze považovat například:

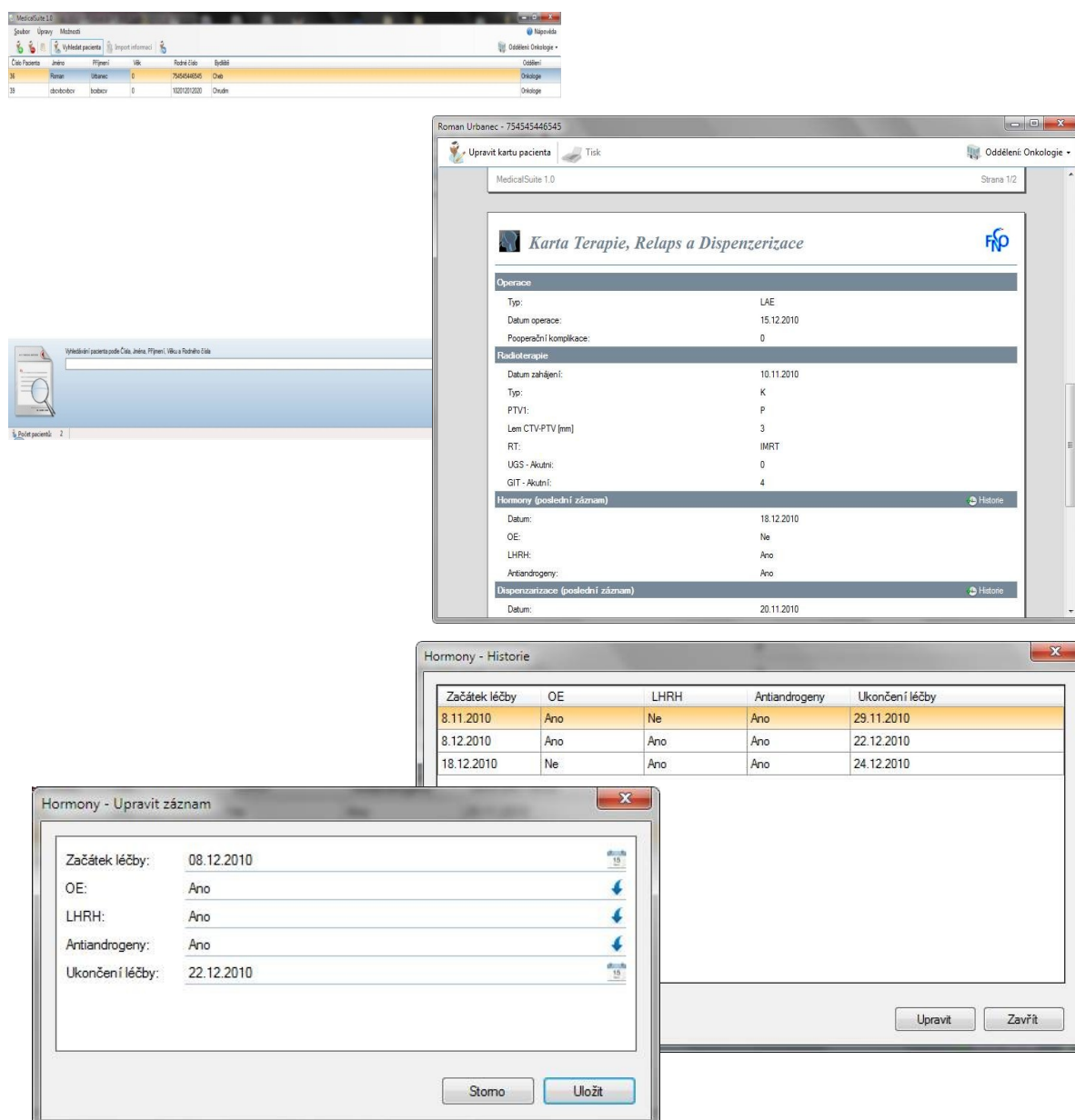
- nesjednocený formát dat (numerická proměnná bývá uváděna jak v číselném formátu, tak v podobě textového řetězce),
- nesjednocené značení variant proměnných (např. chybějící údaje bývají označeny prázdným polem, stejně jako variantami „nelze“, „X“, „neuvedeno“, ...),
- stejný význam různých označení variant proměnných (např. kombinace několika léčebných metod může být značena variantami „RAPE, OE“, „RAPE+OE“, „RAPE a OE“).

3.1 SOFTWARE MEDICALSUITE

Prvním cílem předkládané dizertační práce byla organizace sběru dat spojených s léčbou karcinomu prostaty. Na základě studia dané problematiky a díky mnoha konzultacím s MUDr. Ing. Jakubem Cvekem, PhD. (FNO, Klinika onkologická) byl v rámci grantu SGA FEI VŠB-TUO (projekt

číslo SP/201070) navržen a realizován software MedicalSuite, který lékařům umožňuje záznam relevantních údajů pro následné statistické zpracování.

Software MedicalSuite (viz [Obr. 3.1](#)) byl vytvořen v uživatelsky příjemném prostředí dle požadavků FNO. Slouží k vytváření a shromažďování lékařských záznamů o pacientech, o průběhu jejich hormonální léčby a průběhu a následcích radioterapie. Dále umožňuje import relevantních údajů zaznamenaných v jiných, ve FNO používaných, programech a export záznamů ve formátu vhodném pro následné statistické vyhodnocení. Databáze software MedicalSuite umožňuje vkládání řádově stovek tisíc pacientů, přičemž každý pacient v databázi má cca 50 standardních hodnot a neomezený počet hodnot přídatných.



Obr. 3.1 : Ukázka náhledů na jednotlivé karty fiktivního pacienta zavedeného do MedicalSuite

4 INTERVALOVÉ ODHADY INCIDENCE JEDNOTLIVÝCH TYPŮ TOXICITY

Incidence jednotlivých stupňů toxicity příslušných typů u pacientů z naší studie je uvedena v [Tab. 4.1](#).

Tab. 4.1 : Incidence jednotlivých stupňů toxicity u pacientů z analyzované studie

Typ nežádoucích účinků	Typ toxicity					Celkem
	0	1	2	3	4	
UGS akutní	159	189	52	6	2	408
	39%	46%	13%	1%	0%	
UGS pozdní	119	34	11	6	0	170
	70%	20%	6%	4%	0%	
GIT akutní	234	131	40	3	0	408
	57%	32%	10%	1%	0%	
GIT pozdní	133	23	8	4	0	168
	79%	14%	5%	2%	0%	

V biomedicínských aplikacích se, podobně jako v tomto případě, velmi často setkáváme s poměrně nízkou incidencí některých kategorií sledované proměnné, resp. s malým rozsahem výběru. To jsou situace, které z důvodu porušení předpokladů plynoucích z centrální limitní věty kontraindikují možnost použití standardně využívaného Waldova intervalu spolehlivosti pro kvantifikaci nejistot v podobě intervalového odhadu relativní četnosti (např. incidence). Dalším z cílů dizertační práce bylo nalézt vhodnou metodu pro odhad nejistot při určování incidence sledovaných jevů, vhodnou pro využití v biomedicínských aplikacích. V literatuře (například: [1], [2], [3], [4], [5], [7], [9], [10]) lze najít zhruba 20 různých alternativních metod umožňujících stanovit intervalový odhad parametru binomického rozdělení (angl. „confidence interval for binomial proportion“). V předkládané dizertační práci je uveden přehled vybraných metod pro vyčíslení nejistot formou intervalových odhadů a statistik používaných k hodnocení jejich vlastností, metody jsou porovnány a pomocí vybrané metody jsou následně stanoveny 95% intervalové odhady relativního zastoupení (incidence) jednotlivých stupňů akutní a pozdní toxicity UGS a GIT.

Pro posouzení vhodnosti použití analyzovaných typů intervalových odhadů (dále IO) byly použity statistiky umožňující charakterizovat přednosti i nedostatky těchto odhadů.

Pravděpodobnost pokrytí $C(n, \pi)$ (angl. „coverage probability“) parametru π binomického rozdělení $Bi(n; \pi)$ je definována jako

$$C(n, \pi) = P(\pi \in IO) = \sum_{x=0}^n I_x(\pi, x) \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad (4.1)$$

kde $I_x(\pi, x) = \begin{cases} 1, & \pi \in \langle \pi_D(x, n), \pi_H(x, n) \rangle \\ 0, & \pi \notin \langle \pi_D(x, n), \pi_H(x, n) \rangle \end{cases}$ je tzv. indikátor (angl. „coverage flag“).

Přestože očekáváme, že pravděpodobnost pokrytí parametru π binomického rozdělení $Bi(n; \pi)$ bude pro všechny používané IO blízka specifikované úrovní spolehlivosti $1 - \alpha$ (tzv. nominální pravděpodobnosti pokrytí), realita je, jak bude ukázáno, jiná. Z tohoto pohledu rozlišujeme dva typy intervalů. Intervaly, jejichž minimální pravděpodobnost pokrytí ($\min\{C(n, \pi) | n \in \mathbb{N}, \pi \in \langle 0; 1 \rangle\}$) je menší než nominální pravděpodobnost pokrytí $1 - \alpha$ jsou nazývány **liberální** a intervaly, jejichž minimální pravděpodobnost pokrytí je větší než nominální se nazývají **konzervativní**.

Má-li pravděpodobnost π rovnoměrné rozdělení na intervalu $\langle 0; 1 \rangle$, pak je střední hodnota pravděpodobnosti pokrytí, tzv. **střední pokrytí**, definována jako

$$MC(n) = \int_0^1 C(n, \pi) d\pi \quad (4.2)$$

a **střední kvadratická chyba pokrytí** je definována vztahem

$$RMSE(n) = \sqrt{\int_0^1 (C(n, \pi) - (1 - \alpha))^2 d\pi}. \quad (4.3)$$

Dalším důležitým měřítkem chování intervalového odhadu je jeho délka. Pro hodnocení kvality odhadu je používána délka intervalového odhadu (angl. „expected length“) a střední délka intervalového odhadu (angl. „mean expected length“). **Délka intervalového odhadu** parametru π binomického rozdělení $Bi(n; \pi)$ je

$$EL_{n,\pi}(\text{délka IO}) = \sum_{x=0}^n (\pi_D(x, n) - \pi_H(x, n)) \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad (4.4)$$

kde $\pi_D(x, n)$ a $\pi_H(x, n)$ jsou dolní a horní mez příslušného intervalového odhadu. **Střední délku intervalového odhadu** pak vypočteme jako

$$MEL(n) = \int_0^1 EL_{n,\pi}(\text{délka IO}) d\pi. \quad (4.5)$$

4.1 POROVNÁNÍ ANALYZOVANÝCH INTERVALOVÝCH ODHADŮ

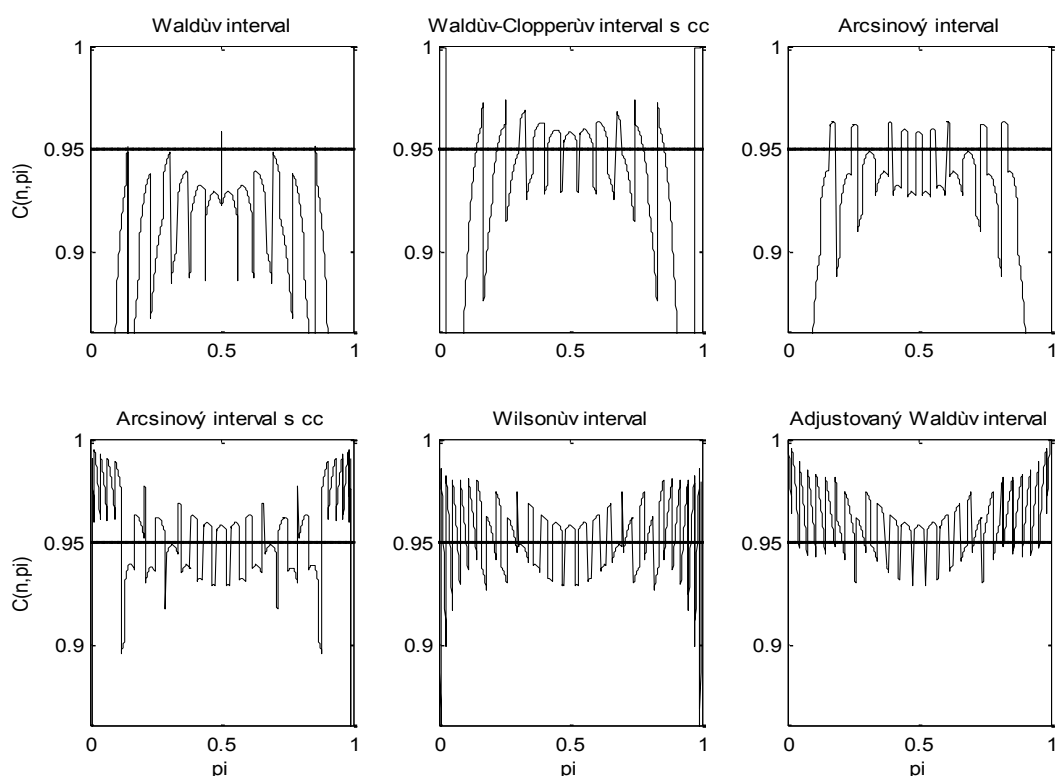
Komplexnímu srovnání analyzovaných liberálních i konzervativních intervalových odhadů parametru π binomického rozdělení, jejichž přehled je uveden v předkládané dizertační práci, je věnována tato kapitola.

4.1.1 POROVNÁNÍ ANALYZOVANÝCH LIBERÁLNÍCH INTERVALOVÝCH ODHADŮ

Analýzy pravděpodobnosti pokrytí byly zaměřeny na biomedicínské aplikace, v nichž se často setkáváme s velmi nízkou incidencí jednotlivých tříd kategoriálních proměnných (např. incidence vysokých stupňů nežádoucích účinků radioterapie). Z analýzy Waldova intervalového odhadu je zřejmé, že tento interval nedává vždy dobré výsledky. Podle obecně uváděných doporučení by měl být Waldův interval používán pouze v případech, kdy rozsah výběru

- je dostatečně velký ($n > 30$),
- je menší než 5% rozsahu základního souboru ($n < 0,05N$),
- splňuje podmínku $n \cdot \min(p, 1 - p) > 9$.

Zvláště splnění podmínky $n \cdot \min(p, 1 - p) > 9$ je v biomedicínských aplikacích, v nichž se často setkáváme s nízkou incidencí sledovaných jevů, obtížné dosáhnout. Pohybuje-li se hledaná pravděpodobnost kolem hodnoty 0,03, tj. 3%, měli bychom pro nalezení intervalového odhadu této pravděpodobnosti s využitím Waldova intervalu mít k dispozici výběr o rozsahu alespoň 310 jednotek. Pro odhad hledané pravděpodobnosti pohybující se kolem hodnoty 0,02, resp. 0,01, by pak byl žádoucí výběr o rozsahu 460, resp. 910 jednotek.

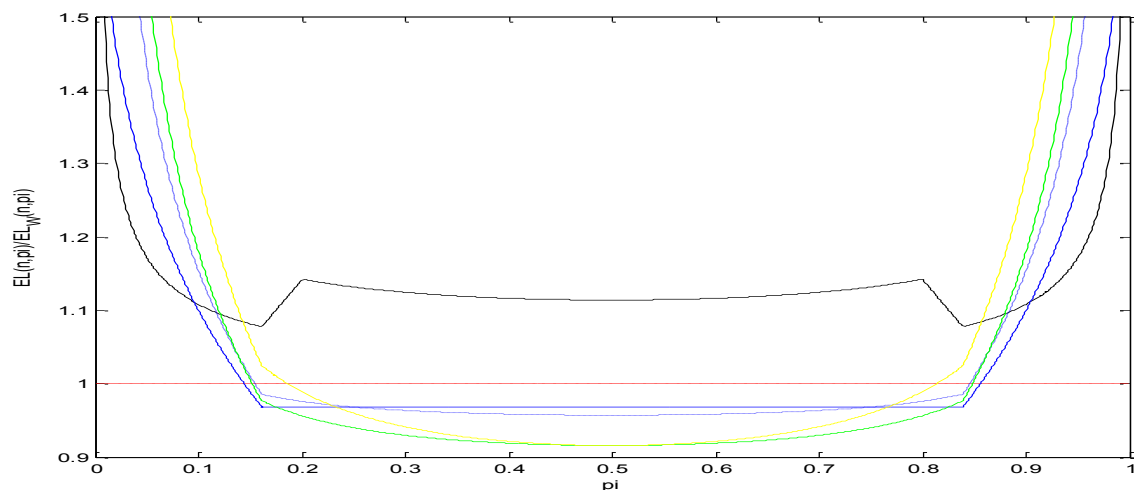


Obr. 4.1: Srovnání pokrytí analyzovaných liberálních intervalů pro $n = 20$, $1 - \alpha = 0,95$

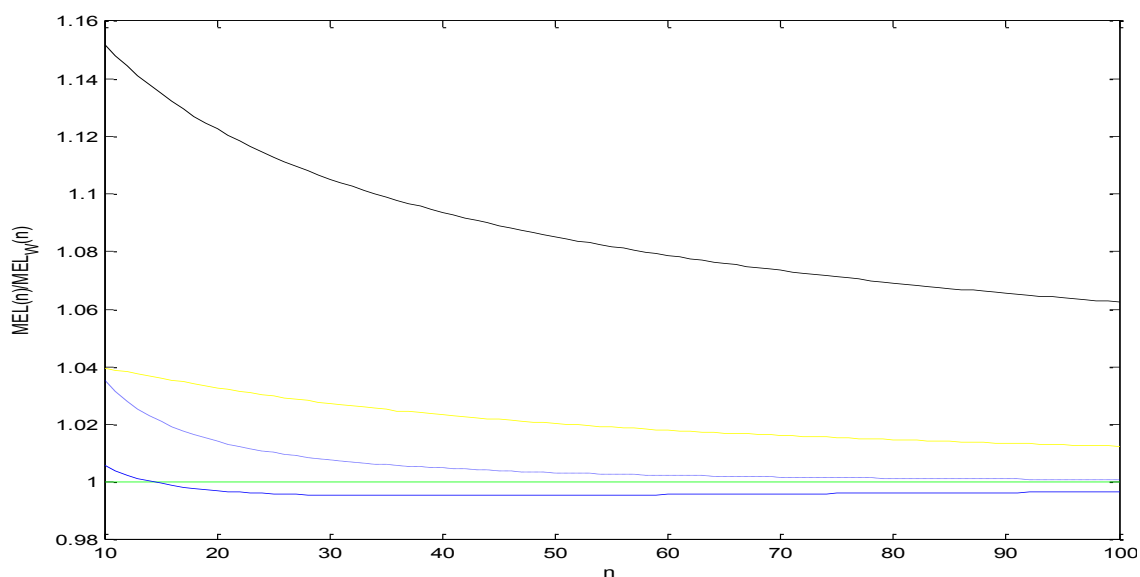
Srovnáním pravděpodobnosti pokrytí Waldova intervalu a ostatních analyzovaných liberálních intervalových odhadů lze dojít k následujícím závěrům (viz [Obr. 4.1](#)): O něco lepší výsledky než Waldův interval poskytuje interval arcsinový, popř. Waldův-Clopperův interval s cc. Přiblížení se pravděpodobnosti pokrytí k nominální hodnotě $1 - \alpha$ je u těchto intervalů rychlejší, počet „šťastných“ dvojic $(n; \pi)$ je vyšší. Posuzujeme-li analyzované intervaly pouze z hlediska pokrytí, jeví se jako vhodný Wilsonův interval, jehož pravděpodobnost pokrytí osciluje kolem nominální hodnoty $1 - \alpha$ pro všechna π na intervalu $\langle 0; 1 \rangle$, arcsinový interval II zajišťující vyšší pokrytí pro hodnoty π blízké 0 nebo 1, resp. adjustovaný Waldův interval vykazující velmi přijatelné minimum pravděpodobnosti pokrytí.

Srovnáním průběhů pravděpodobnosti pokrytí analyzovaných liberálních intervalových odhadů pro $\pi = 0,01$, $1 - \alpha = 0,95$ a $10 \leq n \leq 1000$ dojdeme k obdobným závěrům jako v případě analýzy pravděpodobnosti pokrytí v závislosti na skutečné pravděpodobnosti π . Nejlepší pokrytí (a to i pro malé rozsahy výběru) lze očekávat u arcsinového intervalu II, Wilsonova intervalu a adjustovaného Waldova intervalu.

Pro srovnání analyzovaných liberálních odhadů z hlediska jejich délky bylo použito srovnání poměrů očekávané délky analyzovaných liberálních odhadů k délce Waldova intervalu pro $n = 20$, $1 - \alpha = 0,95$ (viz [Obr. 4.2](#)) a porovnání poměrů střední délky jednotlivých IO a střední délky Waldova IO v závislosti na n (viz [Obr. 4.3](#)). Nejmenší střední délku ze všech analyzovaných liberálních IO vykazuje pro $n > 12$ arcsinový interval. Jeho pokrytí však není optimální. Preferujeme-li u odhadu jeho minimální délku, lze doporučit arcsinový interval II, který nepřekračuje střední délku Waldova IO o více než 4%. Pro $n > 25$ klesne rozdíl mezi střední délkou Waldova IO a arcsinového intervalu II na méně než 1%. Ještě lepších výsledků dosahuje Wilsonův interval, který vykazuje stejnou střední délku jako Waldův interval pro všechna $n = 10, 11, \dots, 100$, a jeho pokrytí osciluje kolem nominální hodnoty. Adjustovaný Waldův interval, který má nejvyšší minimální pokrytí z analyzovaných intervalů, lze z hlediska střední délky rovněž považovat za vhodný. Rozdíl mezi střední délkou Waldova IO a adjustovaného Waldova intervalu se pohybuje mezi 4% a 1%. Výrazně největší střední délku z analyzovaných IO má Waldův-Clopperův interval s cc převyšující střední délku IO až o 15%. Z [Obr. 4.2](#) lze usuzovat na to, že v oblasti π blízkých 0 nebo 1 je při daném rozsahu výběru n délka Waldova IO výrazně nejmenší. Druhou nejmenší délku pak lze pozorovat u Waldova-Clopperova IO.



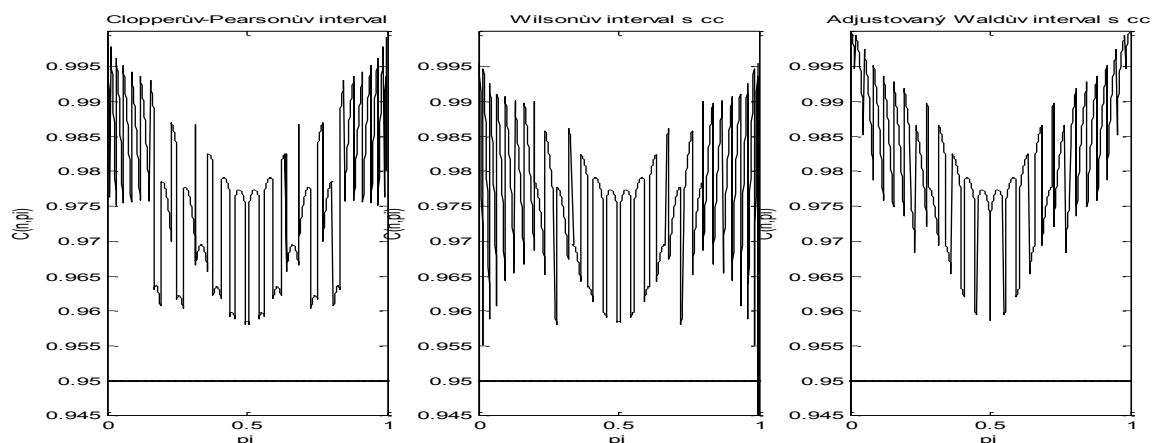
Obr. 4.2: Srovnání poměrů očekávané délky analyzovaných liberálních odhadů k délce Waldova intervalu pro $n = 20$, $1 - \alpha = 0,95$ (Waldův-Clopperův odhad s cc – černá, arcsinový odhad – modrá plná, arcsinový odhad II – modrá čerchovaná, Wilsonův odhad – zelená, adjustovaný Waldův odhad – žlutá)



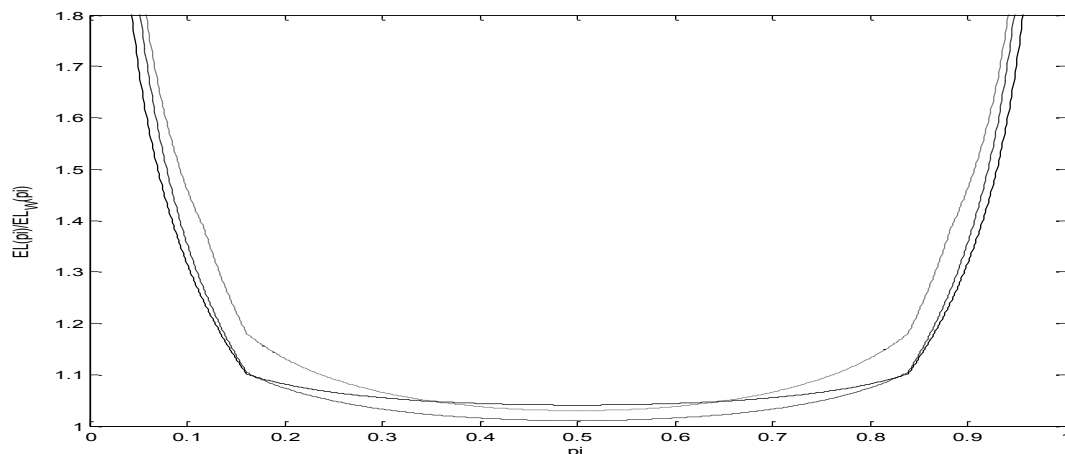
Obr. 4.3: Srovnání poměrů střední délky analyzovaných liberálních odhadů ke střední délce Waldova intervalu v závislosti na n ($1 - \alpha = 0,95$, Waldův-Clopperův odhad s cc – černá, arcsin. odhad – modrá plná, arcsinový odhad II – modrá čerchovaná, Wilsonův odhad – zelená, adjustovaný Waldův odhad – žlutá).

4.1.2 POROVNÁNÍ ANALYZOVANÝCH KONZERVATIVNÍCH INTERVALOVÝCH ODHADŮ

Je-li nutno zajistit konzervativní pokrytí, lze volit mezi Clopperovým-Pearsonovým IO, adjustovaným Waldovým intervalem s cc a Wilsonovým intervalem s cc. Tyto IO garantují minimálně nominální pokrytí parametru binomického rozdělení (tj. jejich pokrytí neklesne pod nominální hodnotu pro žádné π na intervalu $(0; 1)$).



Obr. 4.4: Srovnání pokrytí analyzovaných konzervativních intervalů pro $n = 20$, $1 - \alpha = 0,95$



Obr. 4.5: Srovnání poměrů délky analyzovaných konzervativních odhadů k délce Waldova intervalu pro $n = 20$, $1 - \alpha = 0,95$ (Clopperův-Pearsonův interval – plná, Wilsonův interval s cc – čerchovaná, adjustovaný Waldův interval – tečkovaná)

Pokrytí těchto intervalů lze srovnat na základě Obr. 4.4. Jak lze očekávat, cenou za vysoké hodnoty pokrytí je velká střední délka těchto intervalových odhadů. Zatímco střední délka liberálních odhadů nepřekročila ve většině případů střední délku Waldova intervalu o více než 4%, střední délka konzervativních odhadů překračuje pro rozsah výběru $n = 10$ střední délku Waldova intervalu o 15% až 22%. Při $1 - \alpha = 0,95$ má pro $10 \leq n \leq 32$ nejmenší střední délku Wilsonův odhad s cc. Pro

$n \geq 33$ má nejmenší střední délku Clopperův-Pearsonův interval, rozdíl ve střední délce Clopperova-Pearsonova intervalu a Wilsonova intervalu s cc lze však pro $n \geq 33$ považovat za minimální. Srovnáme-li průběh očekávané délky konzervativních odhadů v závislosti na π (viz Obr. 4.5), lze opět pozorovat srovnatelnou očekávanou délku Clopperova-Pearsonova intervalu a Wilsonova intervalu s cc pro hodnoty π blízké 0 nebo 1. V okolí $\pi = 0,5$ se pak srovnává očekávaná délka všech analyzovaných konzervativních odhadů.

4.2 SOFTWARE PRO VÝPOČET INTERVALOVÉHO ODHADU PARAMETRU BINOMICKÉHO ROZDĚLENÍ

Požadavek na odhad parametru binomického rozdělení se v biomedicínské praxi vyskytuje velmi často. Srovnáme-li ve statistické praxi nejčastěji používané programy (SAS, S-PLUS, SPSS, STATGRAPHICS, R), dojdeme k následujícím závěrům. SAS 9.2 umožňuje výpočet Waldova, adjustovaného Waldova, exaktního Clopperova-Pearsonova, Jeffreysova a Wilsonova skórového intervalu, přičemž u Clopperova-Pearsonova a Jeffreysova intervalu je dolní mez IO nastavena pro $x = 0$ na nulu a pro $x = n$ na jedničku. S-PLUS 8 poskytuje Waldův, exaktní Clopperův-Pearsonův a Wilsonův IO (i s korekci na spojitost). SPSS až do verze 17 neumožňoval výpočet intervalových odhadů parametru binomického rozdělení. SPSS 18 nabízí Waldův, Jeffreysoův a Clopperův-Pearsonův IO. Statgraphics Centurion používá exaktní Clopperův-Pearsonův IO a pro volně šiřitelný software R byla v současné době vytvořena makra pro výpočet osmi různých IO parametru binom. rozdělení (Dorai-Raj **Chyba! Nenalezen zdroj odkazů.**). Vzhledem k tomu, že lékaři obvykle nemají pro základní analýzu svých dat k dispozici statistický software, byl pro usnadnění aplikace těchto intervalů v praxi navržen výpočetní applet IO_binom.xlsx (Obr. 4.6), který umožňuje automaticky výpočet všech intervalových odhadů analyzovaných v předkládané dizertační práci. Pro návrh appletu byl použit běžně dostupný software MS Excel, přičemž pro využití appletu lze využít rovněž volně dostupné OpenOffice.

	A	B	C	D	E	F	G	H	I	J
12										
13			rozsah výběru n :	34						
14			počet úspěchů x :	2						
15			spolehlivost odhadu 1- α :	0,95						
16										
17			Výsledky:							
18										
19			Bodový odhad parametru π :	0,0588						
20										
21			Liberalní intervalové odhady							
22			Název	Odhad se spolehlivostí 0,95	Délka odhadu					
23			Adjustovaný Waldův interval	<0,0065; 0,2007>	0,1942					
24			Wilsonův interval	<0,0163; 0,1909>	0,1747					
25			Arcsinový interval s cc	<0,0143; 0,192>	0,1777					
26			Waldův-Clopperův interval s cc	<0; 0,1379>	0,1379					
27			Arcsinový interval	<0,0059; 0,1611>	0,1552					
28			Waldův interval	<0; 0,1379>	0,1379					Minimální požadovaný rozsah výběru pro použití Waldova intervalu je 163.
29										
30			Konzervativní intervalové odhady							
31			Název	Odhad se spolehlivostí 0,95	Délka odhadu					
32			Clopperův-Pearsonův interval	<0,0072; 0,1968>	0,1896					
33			Wilsonův interval s cc	<0,0103; 0,2106>	0,2003					
34			Adjustovaný Waldův interval s cc	<0,0065; 0,2007>	0,1942					
35										
36										

Obr. 4.6: Náhled na výpočetní applet IO_binom.xlsx - výsledky

5 IDENTIFIKACE Vlivu RADIOTERAPIE A DALŠÍCH FAKTORŮ NA INTENZITU NEŽ. ÚČINKŮ RADIOTERAPIE

Pro analýzu vlivu radioterapie, resp. dalších faktorů (výskyt diabetu, klinické stádium karcinomu prostaty, ...) na intenzitu nežádoucích účinků je možno využít korelační analýzu. Pro korelační analýzu biomedicínských dat je vhodné používat specifické, zatím málo známé, biseriální, polyseriální a polychorické korelace. **Biseriální korelace** vyskytující se ve čtyřech různých modifikacích (bodová, biseriální, modifikovaná biseriální a pořadová biseriální korelace) popisují korelaci mezi dichotomickou a spojitou proměnnou. **Polychorické korelace** lze považovat za zobecnění korelace biseriální.

Mějme skrytou spojitou náhodnou veličinu Z . Necht' X je diskrétní náhodná veličina, pro níž platí:

$$X = x_j \text{ pokud } \tau_{j-1} \leq Z \leq \tau_j, \quad j = 1, 2, \dots, s, \quad (5.1)$$

kde s je počet variant náhodné veličiny X a τ_j jsou hranice. Necht' hranice $\tau_0 = -\infty, \tau_s = \infty$ a necht' jsou hranice, stejně jako diskrétní náhodná veličina X , ostře rostoucí, tj. $\tau_0 < \tau_1 < \tau_2 < \dots < \tau_s$, $x_1 < x_2 < \dots < x_s$. Předpokládejme, že (Z, Y) má dvourozměrné normální rozdělení s parametry $E(Y) = \mu$, $D(Y) = \sigma^2$, $E(Z) = 0$, $D(Z) = 1$ a $\frac{\text{cov}(Z, Y)}{\sigma} = \rho$. Polyseriální korelace ρ_p mezi diskrétní náhodnou veličinou X a spojitou náhodnou veličinou Y je definována jako korelace ρ mezi Y a skrytou normálně rozdělenou veličinou Z vztahující se k X .

Pro měření intenzity závislosti mezi dvěma diskrétními proměnnými lze použít Spearmanův korelační koeficient, popřípadě koeficienty pro měření intenzity závislosti v kontingenční tabulce. V biomedicínských aplikacích se často setkáváme s myšlenkou, že diskrétní proměnné jsou odvozeny kategorizací nějakých skrytých spojitých proměnných. To je hlavní myšlenkou **polychorických korelací**, které tak jsou další z možností, jak měřit intenzitu závislosti mezi diskrétními proměnnými vzniklými kategorizací skrytých spojitých proměnných, tj. např. korelaci mezi výskytem diabetu u pacienta a intenzitou nežádoucích účinků radioterapie.

5.1 SOFTWARE PRO VÝPOČET KORELAČNÍCH KOEFICIENTŮ

Srovnáme-li ve statistické praxi obvykle používané programy (SAS, SPSS, Statgraphics), dojdeme k závěru, že algoritmy pro výpočet uvedených specifických korelačních koeficientů nejsou běžnou součástí statistického software.

Pro usnadnění aplikace vhodných korelačních koeficientů v praxi byly v rámci předkládané dizertační práce navrženy výpočetní applety [polyserial.xlsm](#), [korelace_diskret.xlsm](#) a

[korelace_spoj.xlsx](#), které umožňují automatický výpočet korelačních koeficientů určených pro analýzu vztahů mezi diskrétní a spojitou proměnnou, popř. mezi dvěma diskrétními proměnnými nebo korelace mezi dvěma spojitými proměnnými. Pro návrh appletů byl použit běžně dostupný software MS Excel, přičemž pro využití appletu lze využít rovněž volně dostupné OpenOffice.

Applet [polyserial.xlsm](#) je určen k analýze závislosti mezi diskrétní a spojitou proměnnou. V případě, že diskrétní proměnná je dichotomická, je prezentována bodová biseriální korelace, biseriální korelace a pořadová biseriální korelace, včetně testu významnosti, a modifikovaná biseriální korelace. Nabývá-li diskrétní proměnná více než dvou variant, je výstupem bodový polyseriální koeficient a polyseriální koeficient vypočtený metodou Ad Hoc a dvoukrokovou metodou. Jako doplňková informace jsou prezentovány základní číselné charakteristiky spojitě proměnné, doplněné testem normality a základní podmíněné číselné charakteristiky spojitě proměnné v závislosti na hodnotě diskrétní proměnné (opět včetně testů normality). Pro testy normality byly použity testy založené na hodnotách šikmosti a špičatosti.

Applet [korelace_diskret.xlsm](#) slouží k výpočtu Spearmanova a Kendallova korelačního koeficientu, přičemž vstupní data mohou být zadána jak ve standardním datovém formátu, tak pomocí tabulky sdružených četností. Při výpočtu Kendallova koeficientu je prováděna automatická volba vhodné varianty (τ_a , τ_b , nebo τ_c). Jako doplňková informace je vždy prezentována tabulka četností, v níž lze zobrazit sdružené a marginální četnosti, sdružené a marginální relativní četnosti, řádkové četnosti, popř. sloupcové četnosti. Výstup z této tabulky pak lze použít jako vstupní data volně šiřitelného programu POLYCORR [7] umožňujícího výpočet polychorické korelace, včetně její směrodatné odchylky. (Součástí vstupu programu POLYCORR je tabulka sdružených četností ve formátu txt.)

Applet [korelace_spoj.xlsx](#) slouží k výpočtu Pearsonova a Spearmanova korelačního koeficientu. Jako doplňková informace jsou prezentovány rozptylogram a histogramy obou proměnných pro orientační posouzení normality. Při výpočtu Pearsonova korelačního koeficientu je zobrazován výsledek testu normality (test založený na kombinaci šikmosti a špičatosti) obou analyzovaných proměnných.

6 VYHODNOCENÍ INCIDENCE A PREDIKTORŮ TOXICITY PROVÁZEJÍCÍ RADIOTERAPII KARCINOMU PROSTATY

Radioterapie je účinnou metodou pro léčbu celé řady malignit. V souvislosti s vysokou dávkou záření je s radioterapií spojena i toxicita zdravých tkání. Studie analýzy nežádoucích účinků spojených s radioterapií karcinomu prostaty proběhla v letech 2006-2009 ve FNO pod vedením MUDr. Ing. Jakuba Cveka, PhD. Do studie bylo zařazeno 408 pacientů, kteří v souvislosti s léčbou karcinomu

prostaty podstoupili radioterapii a zároveň byli sledováni alespoň 3 měsíce od ukončení radioterapie (rozmezí doby sledování: 3-36 měsíců, medián doby sledování: 8 měsíců). Pacientům byla na oblast GIT aplikovaná maximální dávka v rozmezí (43,5-92,5) Gy, medián maximální dávky byl 72,8 Gy. Na oblast UGS byla aplikovaná maximální dávka v rozmezí (34,6-92,8) Gy, medián maximální dávky byl 71,8 Gy.

Většina pacientů (324/408, tj. 79,2%) neměla žádné nebo pouze mírné (stupeň 1) projevy akutní toxicity. Jako závažnější pro kvalitu života pacienta jsou vnímány projevy pozdní toxicity. Proto lze jako velmi příznivou hodnotit i skutečnost, že obdobně příznivé výsledky byly pozorovány v projevech pozdní toxicity, kdy žádné nebo mírné projevy byly pozorovány u 142/164, tj. 86,6% pacientů, žádný z pacientů neudával projevy pozdní toxicity odpovídající stupni 4. Byli zaznamenáni 3 pacienti udávající projevy pozdní toxicity UGS stupně 3, resp. 4 pacienti udávající projevy pozdní toxicity GIT stupně 3.

Při zachování stávajícího způsobu léčby lze se spolehlivostí 0,95 očekávat, {0,1%; 2,2%} pacientů se závažnými (stupeň 3 nebo 4) projevy pozdní toxicity UGS a {0,7%; 6,2%} pacientů se závažnými projevy pozdní toxicity GIT.

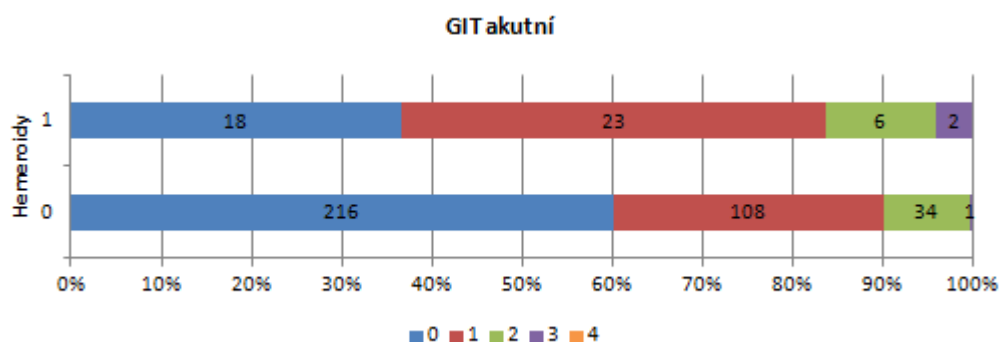
Jako potenciální prediktory toxicity UGS a GIT byly vyhodnocovány: riziková skupina (odpovídající klinickému stádiu karcinomu), warfarizace, onemocnění diabetem melitus a hemeroidy a přítomnost dalšího karcinomu. Riziková skupina, do níž je pacient zařazen se projevovala více na intenzitě projevů akutní toxicity než na projevech toxicity pozdní, korelace mezi rizikovou skupinou a intenzitou nežádoucích účinků však nebyla vyhodnocena jako statisticky významná. Jako statisticky nevýznamné prediktory toxicity UGS a GIT byly vyhodnoceny rovněž onemocnění diabetem, warfarizace a přítomnost dalšího karcinomu. Onemocnění hemeroidy bylo vyhodnoceno jako statisticky významný (na hladině významnosti 0,05) prediktor akutní i pozdní toxicity GIT (viz [Obr. 6.1](#), [Obr. 6.2](#), [Tab. 6.1](#) - [Tab. 6.3](#)).

Tab. 6.1: Kendallovo τ_c a p -hodnota pro test nulovosti τ_c

τ_c (p -hodnota)	Hemeroidy
UGS akutní	-0,0097 (0,7771)
GIT akutní	0,1020 (0,0019)
UGS pozdní	-0,0425 (0,3965)
GIT pozdní	0,1402 (0,0020)

Tab. 6.2: Pozorovaná incidence jednotlivých stupňů akutní toxicity GIT v závislosti na onemocnění hemeroidy

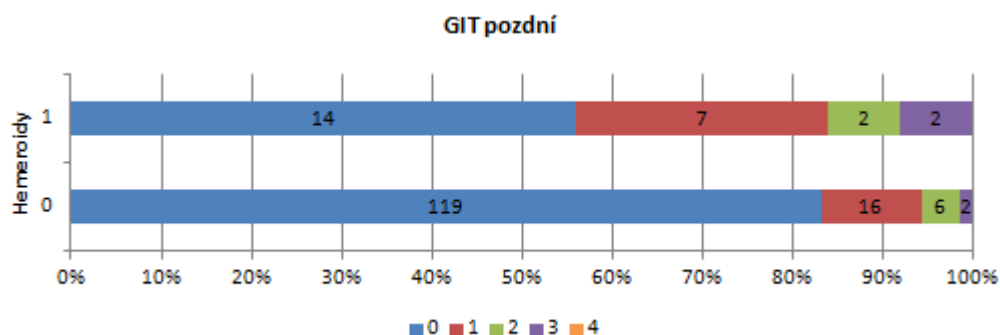
Hemeroidy \ GIT akutní	0	1	2	3	4	Celkem
0	216	108	34	1	0	359
1	18	23	6	2	0	49
Celkem	234	131	40	3	0	408



Obr. 6.1: Závislost incidence jednotlivých stupňů akutní toxicity GIT v závislosti na onemocnění hemeroidy

Tab. 6.3: Pozorovaná incidence jednotlivých stupňů pozdní toxicity GIT v závislosti na onemocnění hemeroidy

Hemeroidy \ GIT pozdní	0	1	2	3	4	Celkem
0	119	16	6	2	0	143
1	14	7	2	2	0	25
Celkem	133	23	8	4	0	168



Obr. 6.2: Závislost incidence jednotlivých stupňů pozdní toxicity GIT v závislosti na onemocnění hemeroidy

Jako statisticky významná byla hodnocena rovněž korelace mezi akutními a pozdními projevy toxicity UGS i GIT. Přetrvávající zvýšená (stupeň >2) toxicita UGS byla pozorována u 8/36, tj. 22,2 % pacientů, přetrvávající zvýšená toxicita GIT byla pozorována u 2/19, tj. 10,5 % pacientů.

Korelace mezi parametry DVH reprezentovanými ekvivalentní uniformní dávkou (EUD) a intenzitou nežádoucích účinků byla téměř nulová (0,023) v případě pozdní toxicity UGS. V ostatních případech byla pozorována středně silná korelace mezi *EUD* a příslušným typem toxicity.

7 ZÁVĚRY PLYNOUCÍ Z PŘEDLOŽENÉ DIZERTAČNÍ PRÁCE

Závěry plynoucí z předložené dizertační práce mohou být rozděleny do dvou oblastí:

- závěry plynoucí pro praxi,
- závěry plynoucí pro rozvoj vědy.

7.1 ZÁVĚRY PLYNOUCÍ PRO PRAXI

Předložená dizertační práce se zabývá aplikací moderních statistických metod v analýze biomedicínských dat. Hlavními cíli dizertační práce byly:

1. Organizace sběru dat spojených s léčbou karcinomu prostaty a generování softwarové aplikace pro jejich předzpracování.
2. Analýza možných metod odhadu nejistot při určování incidence sledovaných jevů vycházející ze simulace pravděpodobnosti pokrytí konfidenčních intervalů. Generování výpočetního appletu pro kvantifikaci a vzájemné porovnání těchto nejistot.
3. Hlubková korelační analýza modifikovaná pro biomedicínská data a jejich specifické zvláštnosti. Generování výpočetního appletu umožňujícího vyčíslení specifických korelačních koeficientů.
4. Tvůrčí vyhodnocení záznamů o radioterapeutické léčbě karcinomu prostaty (2006 – 2009, MUDr. Ing. Jakub Cvek, PhD., Fakultní nemocnice Ostrava, Klinika onkologická), identifikace závislostí radioterapie a dalších faktorů v kontextu s jejími nežádoucími účinky.

Ad1) Na základě studia dané problematiky a díky mnoha konzultacím s MUDr. Ing. Jakubem Cvekem, PhD. (FNO, Klinika onkologická) byl v rámci grantu SGA FEI VŠB-TUO (projekt číslo SP/201070) navržen a realizován software MedicalSuite, který lékařům umožňuje záznam relevantních údajů pro následné statistické zpracování. Software MedicalSuite slouží k vytváření a shromažďování lékařských záznamů o pacientech, o průběhu jejich hormonální léčby a průběhu a následcích radioterapie, umožňuje import relevantních údajů zaznamenaných v jiných ve FNO používaných programech a export záznamů ve formátu vhodném pro následné statistické vyhodnocení.

Ad2) V předkládané dizertační práci je provedeno srovnání v literatuře popsaných intervalových odhadů parametru binomického rozdělení. Srovnání je provedeno s ohledem na možné využití při analýze biomedicínských dat (nízký rozsah výběru, nízká pozorovaná četnost výskytu sledovaného jevu). V případě, že výzkumník nevyžaduje, aby minimální pokrytí parametru binomického rozdělení bylo vyšší než nominální hodnota (většinou 0,95), lze pro praktické použití doporučit Wilsonův nebo adjustovaný Waldův interval, přičemž adjustovaný Waldův interval vykazuje vyšší pokrytí pro hodnoty parametru binomického rozdělení blízké hraničním hodnotám 0, 1. Cenou za vyšší pokrytí je větší délka tohoto intervalu. Je-li požadováno, aby byl interval konzervativní, jeví se jako vhodná alternativy Waldova intervalu Clopperův-Pearsonův interval, resp. adjustovaný Waldův interval, jež vykazují menší střední kvadratickou chybu odhadu než Wilsonův interval s cc.

Pro snadný výpočet analyzovaných intervalových odhadů byl generován volně šiřitelný výpočetní applet [IO_binom.xlsx](#), který je součástí CD přiloženého k předkládané dizertační práci.

Ad3) Při analýze biomedicínských dat se často setkáváme s požadavkem na určení korelace mezi diskrétní a spojitou proměnnou (např. korelace mezi *EUD* a pozdní toxicitou UGS), resp. mezi dvěma diskrétními proměnnými, které vznikly diskretizací dvou skrytých proměnných s dvourozměrným normálním rozdělením (např. korelace mezi akutní a pozdní toxicitou). V těchto případech nelze pro odhad korelace mezi proměnnými použít obecně známé korelační koeficienty (Pearsonův koeficient, Spearmanův koeficient, Kendallov τ). Kapitola 0 předkládané práce popisuje v české literatuře dosud komplexně nepopsané a ve statistické praxi běžně nepoužívané typy korelačních koeficientů – biseriální, polyseriální a polychorický.

Pro usnadnění aplikace vhodných korelačních koeficientů v praxi byly v rámci předkládané dizertační práce generovány výpočetní applety, které umožňují automatický výpočet korelačních koeficientů určených pro analýzu vztahů mezi diskrétní a spojitou proměnnou ([polyserial.xlsm](#)), popř. mezi dvěma diskrétními proměnnými ([korelace_diskret.xlsm](#)) nebo korelace mezi dvěma spojitými proměnnými ([korelace_spoj.xlsx](#)). Součástí appletů jsou testy významnosti jednotlivých korelačních koeficientů, popř. testy normality pro ověření oprávněnosti použití jednotlivých korelačních koeficientů, resp. další vhodné číselné charakteristiky analyzovaných proměnných.

Ad4) Z hlediska časového vývoje nežádoucích účinků radioterapie se rozlišují dva typy reakcí:

a) akutní, které se vyvíjí obvykle již v průběhu léčby, jsou relativně dobře zvládnutelné a ustupují zpravidla do 2-4 týdnů po ukončení léčby,

- b) pozdní, které vznikají v horizontu několika let po ukončení léčby, jsou trvalého charakteru a mohou významně ovlivnit život pacienta.

V souvislosti s radioterapií karcinomu prostaty je obvykle sledována toxicita trávicího traktu (GIT) a toxicita močového traktu (UGS). Míra jak akutních, tak pozdních účinků radioterapie na námi sledované orgány se hodnotí podle mezinárodně uznávaného skórovacího systému. Z uvedených výsledků je zřejmé, že většina pacientů zařazených do analýzy (324/408, tj. 79,2%) neměla žádné nebo pouze mírné (stupeň 1) projevy akutní toxicity. Závažné projevy pozdní toxicity byly pozorovány u 22/164, tj. u 13,4% pacientů, přičemž žádný z pacientů neudával projevy pozdní toxicity odpovídající stupni 4. Při zachování stávajícího způsobu léčby lze se spolehlivostí 0,95 očekávat, {0,1%; 2,2%} pacientů se závažnými (stupeň 3 nebo 4) projevy pozdní toxicity UGS a {0,7%; 6,2%} pacientů se závažnými projevy pozdní toxicity GIT.

V literatuře [6] se uvádí, že míra pozdních nežádoucích účinků je závislá na dávce záření aplikované na daný orgán. Pro dávky aplikované v naší studii byla pozorována velmi slabá korelace mezi ekvivalentní uniformní dávkou (*EUD*) a pozdní toxicitou UGS. Korelaci mezi *EUD* a ostatními typy toxicity lze označit za středně silnou.

Při identifikaci potencionálních prediktorů toxicity UGS a GIT bylo jako statisticky významný (na hladině významnosti 0,05) prediktor akutní i pozdní toxicity GIT identifikováno onemocnění hemeroidy. Jako statisticky významná (na hladině významnosti 0,05) byla hodnocena rovněž korelace mezi akutními a pozdními projevy toxicity UGS i GIT.

7.2 ZÁVĚRY PLYNOUCÍ PRO ROZVOJ VĚDY

K problematice intervalových odhadů parametru binomického rozdělení, stejně jako k problematice korelací diskrétní a spojitě proměnné existuje sice poměrně široké spektrum zahraniční literatury, avšak ve statistické praxi stále převládá používání Waldova intervalu pro odhad parametru binomického rozdělení, resp. využití Spearmanova koeficientu nebo Kendallova τ pro odhad korelace mezi diskrétní a spojitou proměnnou.

Předkládaná dizertační práce samozřejmě není kompletní studií uvedené problematiky, lze ji však považovat za vodítko při odhadování incidence řídkých jevů, resp. při odhadování specifických korelačních koeficientů. Výpočetní applety, které byly v rámci dizertační práce generovány, mají za cíl usnadnit zavedení vhodných typů intervalových odhadů parametru binomického rozdělení a specifických typů korelací do praxe.

7.3 MOŽNOSTI DALŠÍHO POKRAČOVÁNÍ V ŘEŠENÍ PROBLEMATIKY

Z hlediska dalšího rozvoje problematiky řešené v práci je možno zmínit několik idejí, které by bylo v rámci dalšího řešení vhodné realizovat. Namátkově lze uvést:

- Popis metod pro intervalové odhady rozdílů parametrů dvou binomických rozdělení a jejich algoritmizace.
- Tvorba online výpočetních appletů pro odhad a analýzu nejistot parametru binomického rozdělení.
- Tvorba online výpočetních appletů pro analýzu závislostí specifických biomedicínských proměnných, např. pro odhad specifických korelačních koeficientů, včetně testu jejich nulovosti.

SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK

Značka	Význam
$C(n, \pi)$	pravděpodobnost pokrytí
$El n, \pi(\text{délka IO})$	očekávaná délka intervalového odhadu
EUD	ekvivalentní uniformní dávka [Gy]
FNO	Fakultní nemocnice v Ostravě
GIT	trávicí trakt
GS	Gleasonovo skóre (stupeň diferenciacie nádoru)
I_x	indikátor
$MC(n)$	střední pokrytí
$MEL(n)$	střední délka intervalového odhadu
PSA	koncentrace prostatického antigenu [ng/ml]
RAPE	radikální prostatektomie
$RMSE(n)$	střední kvadratická chyba pokrytí
UGS	močový trakt
cc	korekce na spojitost
n	rozsah výběru
p	bodový odhad pravděpodobnosti π $\left(= \frac{x}{n}, \text{ kde } x \text{ je počet výskytů sledované události a } n \text{ je rozsah výběru} \right)$
α	hladina významnosti
π	skutečná pravděpodobnost výskytu sledovaného jevu
π_D	dolní mez intervalového odhadu pravděpodobnosti π
π_H	horní mez intervalového odhadu pravděpodobnosti π
ρ	korelační koeficient
τ_i	hranice pro kategorizaci spojitě náhodné veličiny

LITERATURA

- [1] AGRESTI, A., & COULL, B.A. (1998), Approximate is better than “exact” for interval estimation of binomial proportions, *The American Statistician*, Vol. 52, pg. 119–126..
- [2] AGRESTI, A., & CAFFO, B. (2000) Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures, *The American Statistician*, Vol. 54, pg. 280–288.
- [3] ANDĚL, M., ČERNÝ, R., CHARAMZA, P., NEUSTADT, J. (2004), Přehled metod odhadu statistické chyby ve výběrových šetřeních, *Informační bulletin České Statistické Společnosti*, číslo 2-3, ročník 15, ISSN 1210 – 8022.
- [4] BLYTH, C.R., & STILL, H.A. (1983), Binomial confidence intervals, *Journal of the American Statistical Association*, Vol. 78, pg. 108–116.
- [5] CLOPPER, C.J., & PEARSON, E.S. (1934), The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika*, Vol. 26, pg. 404–413.
- [6] KUBEŠ, J. (2010), Moderní trendy v radioterapii karcinomu prostaty, *Onkologie*, Vol. 4(2), pg. 79–83, [cit. 2011-06-27], dostupný na Word Wide Web: <http://www.onkologiecs.cz/pdfs/xon/2010/02/06.pdf>.
- [7] PIRES, A. M., CONCEICAO A. (2008), Interval estimators for binomial proportion: Comparison of twenty methods, *Statistical Journal*, Vol. 6, pg. 165-197.
- [8] UEBERSAX, JS (2007), *User Guide for POLYCORR 1.1*, Statistical Methods for Rater Agreement web site, [cit. 2011-09-26], dostupný na Word Wide Web: <http://john-uebersax.com/stat/xpc.htm>.
- [9] WALD, A., & WOLFOWITZ, J. (1939), Confidence limits for continuous distribution functions, *The Annals of Mathematical Statistics*, Vol. 10, pg. 105–118.
- [10] WILSON, E.B. (1927), Probable inference, the law of succession, and statistical inference, *Journal of the American Statistical Association*, Vol. 22, pg. 209–212.

CURICULUM VITAE

Jméno: Ing. Martina Litschmannová

VZDĚLÁNÍ

- | | |
|-------------|---|
| od 2007 | doktorské studium, VŠB – Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky, studijní obor: Technická kybernetika, název dizertační práce: Využití moderních statistických metod pro analýzu nežádoucích účinků spojených s radioterapií karcinomu prostaty, školitel: prof. Ing. Radim Briš, CSc. |
| 1991 - 1996 | Ing., ČVUT Praha, Fakulta elektrotechnická, Katedra mikroelektroniky, studijní obor: Mikroelektronika, název diplomové práce: Modelování spínaných zdrojů |
| 1987 - 1991 | Gymnázium Thälmannova 669, Ostrava - Poruba |

PRAXE

- | | |
|---------|---|
| od 1996 | Odborná asistentka, VŠB – Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky, Katedra aplikované matematiky
(2000 – 2006) mateřská dovolená |
|---------|---|

ODBORNÁ ČINNOST DOKTORANDKY

ODBORNÁ ČINNOST S PŘÍMÝM VZTAHEM K DIZERTAČNÍ PRÁCI

ODBORNÉ PUBLIKACE NA WEB OF SCIENCE

- [1] BRIŠ R., LITSCHMANNOVÁ M. (2000), Simulation Approach in Reliability Analysis of Maintained Systems, *Proceedings of the European Conference on Safety and Reliability – ESREL 2000*, Edinburgh, Scotland UK, Published by A.A.Balkema 2000, Vol.2, pg. 1563-1569, Rotterdam, ISBN 9058091406.
- [2] LITSCHMANNOVÁ, M. (2008), The methodology for the evaluation of energy savings' states of the public lighting, *In Przegląd Elektrotechniczny*, Vol. 8/08, pg. 65-67, (Impact Factor 0.196).

DALŠÍ ODBORNÉ PUBLIKACE

- [1] LITSCHMANNOVÁ, M., ZDRÁHALA, A. (2010), Side Effects of Prostate Cancer Radiotherapy, *In Sborník konference WOFEX 10*, pg. 320-325, ISBN 978-80-248-2276-1.
- [2] LITSCHMANNOVÁ, M., ZDRÁHALA, A. (2010), Radioterapie karcinomu prostaty a její nežádoucí účinky na okolní tkáň, *In Sborník konference Request 10*.
- [3] LITSCHMANNOVÁ, M., VAŠÍČKOVÁ, H. (2009), Poznámky k analýze medicínských dat, *In Sborník konference Request 09*, ISBN 978-80-7372-619-5, s. 108-113.

- [4] LITSCHMANNOVÁ, M. (2009), Discriminant Analysis of Medical Data , *In Sborník konference Faculty of Natural Sciences University of SS Cyril and Methodius in Trnava, Slovak Republic* , ISBN 978-80-8105-129-6 .
- [5] LITSCHMANNOVÁ, M. (2009), The Modelling of Post-Surgery Morbidity of Patients Facing Colectomy, *In Sborník konference WOFEX 09*, ISBN 978-80-248-2028-6.

PŘEDNÁŠKY

- [1] LITSCHMANNOVÁ, M. (2011), *Waldův intervalový odhad parametru binomického rozdělení a jeho alternativy*, STAKAN & TEXPRIENCE 2011

REGISTROVANÝ SOFTWARE

- [1] LITSCHMANNOVÁ, M., HAVLÍK, L. (2010): *Software MedicalSuite*, ev. č.: 063/14-12-2010_SW.

ÚČAST NA GRANTECH

- [1] Projekt MŠMT ČR programu Výzkumná centra (č. 1M06047) s názvem *Centrum pro jakost a spolehlivost výroby*. Celostátní projekt, hlavní zodpovědný řešitel za VŠB TU Ostrava -prof.Ing. R.Briš, CSc. 2006-2012.
- [2] VONDRÁK, I.: Výzkumný projekt Moravskoslezského kraje 01562/2006/RER „*FLOREON - FLloods REcogniton On the Net*“, financováno z rozpočtu Moravskoslezského kraje, 2007 -2012.
- [3] JAHODA, P., LITSCHMANNOVÁ, M., PRAKS, P.: *Modelování a kvantifikace rizik v lékařství, zpracování biomedicínských dat*, grant SGS SP2011/100.
- [4] JAHODA, P., LITSCHMANNOVÁ, M., PRAKS, P.: *Modelování a kvantifikace rizik v lékařství, zpracování biomedicínských dat*, grant IGA 25/2009 – 2010.

OSTATNÍ ODBORNÁ ČINNOST

ODBORNÉ PUBLIKACE NA WEB OF SCIENCE

- [1] VALA, R., VALOVÁ, M., LITSCHMANNOVÁ, M., KLIMTOVÁ, H. (2010), Sprinting Abilities of Year Six Students Undergoing Additional Physical Education Classes, *The New Educational Review*, Vol. 22, pg. 165-177, ISSN 1732-6729, (Impact Factor 0.040).

ÚČAST NA GRANTECH

- [1] ŠARMANOVÁ, P., BEREMLIJSKI, P., *Matematika pro 21. století*, grant ESF (CZ.1.07/2.2.00/07.0332), **2009-2012** (pozice: garant modulu, statistik, animátor)
- [2] SOKANSKÝ, K., DOSTÁL, F., NOVÁK, T.: *Potenciál energetických úspor veřejného osvětlení v ČR*. MPO: 222004 7305, **2007**
- [3] ŠARMANOVÁ, J., E-learningové prvky pro podporu výuky odborných a technických předmětů, grant ESF (CZ.O4.01.3/3.2.15.2/0326), **2005-2007**

ODBORNÉ PUBLIKACE

- [1] VALA, R., M., LITSCHMANNOVÁ, VALOVÁ, (2011), Úroveň rychlostních schopností děvčat šestých tříd na základní škole s rozšířenou výukou tělesné výchovy, *Studia sportiva*, roč. 5, sv. 1, s. 125-131, ISBN 1802-7679.
- [2] VALA, R., VALOVÁ, M., LITSCHMANNOVÁ, M. (2010), Úroveň rychlostních schopností žáků šestých tříd na základní škole s rozšířenou tělesnou výchovou, *Studia Kinanthropologica*, XI., str. 30-35, ISSN – 1213-2101.

- [3] LITSCHMANNOVÁ, M. (2008), Bayesian Networks as an Instrument used for Processing of Indeterminacy burdened Knowledge, *In Sborník konference WOFEX 08*.
- [4] KLÍMKOVÁ M. (1998), JMP-IN - program, that makes statistics visual, *Proceedings of the scientific conference with international participation - Informatics and Algorithms 1998*, Prešov, Slovakia.
- [5] KLÍMKOVÁ M. (1998), Analog Behavioral Modeling, *Proceedings of the XXth International Workshop - Advanced Simulation of Systems (ASIS) 1998*, Ostrava.
- [6] BRIŠ R., DUMMER R. M., KLÍMKOVÁ M. (1997), Softwarová podpora výuky statistiky na VŠB-TU Ostrava, Sborník přednášek a programů - Pedagogický software 97, Scientific Pedagogical Publishing, České Budějovice.

SKRIPTA

- [1] LITSCHMANNOVÁ, M. (2011), *Úvod do statistiky*, skripta - pilotní verze , dostupné z: <http://mi21.vsb.cz/modul/uvod-do-statistiky>
- [2] LITSCHMANNOVÁ, M. (2011): *Vybrané kapitoly z pravděpodobnosti*, skripta - pilotní verze, dostupné z: <http://mi21.vsb.cz/modul/vybrane-kapitoly-z-pravdepodobnosti>
- [3] LITSCHMANNOVÁ, M. (2008): *Statistika I. – neřešené příklady*, dostupné z: <http://home1.vsb.cz/~lit40/STA1/statistika.html?butt1=Statistika+I>
- [4] LITSCHMANNOVÁ, M. (2007), *Statistika I. - řešené příklady*, dostupné z: <http://home1.vsb.cz/~lit40/STA1/statistika.html?butt1=Statistika+I>
- [5] BRIŠ R., LITSCHMANNOVÁ M. (2007), *Statistika II.*, E-learningový prvek pro podporu výuky odborných a technických předmětů, v rámci projektu CZ.O4.01.3/3.2.15.2/0326, VŠB TU Ostrava, 2007, ISBN 978-80-248-1482-7.
- [6] BRIŠ R., LITSCHMANNOVÁ M. (2004), *Statistika I. pro kombinované a distanční studium*, Ostrava.
- [7] DUMMER R. M., KLÍMKOVÁ M. (1997), *Statistika I. (cvičení)*, skripta VŠB-TUO, Ostrava.

PŘEDNÁŠKY

- [1] LITSCHMANNOVÁ, M. (2011), *Tvorba internetových učebnic pravděpodobnosti a statistiky jako součást projektu MI21*, STAKAN & TEXPRIENCE 2011.
- [2] LITSCHMANNOVÁ, M. (2010), *Bayesův teorém - cesta k lepší náladě*, ŠKOMAM 2010, dostupné z: <http://home1.vsb.cz/~lit40/hlavni/omne.html?butt1=O+mn%EC>
- [3] LITSCHMANNOVÁ, M. (2009), *Máme dotazníky. A co dál?*, ŠKOMAM '09, dostupné z: <http://home1.vsb.cz/~lit40/hlavni/omne.html?butt1=O+mn%EC>
- [4] LITSCHMANNOVÁ, M. (2008), *Na co ve výuce statistiky není čas aneb Pane, pojděte si hrát*, ŠKOMAM '08, dostupné z: <http://home1.vsb.cz/~lit40/hlavni/omne.html?butt1=O+mn%EC>

PREZENTACE NA KONFERENCÍCH

- [1] VALA, R., VALOVÁ, M., LITSCHMANNOVÁ, M. *Prezentace: Úroveň rychlostních schopností děvčat šestých tříd na základní škole s rozšířenou výukou tělesné výchovy*. Hradec Králové: Univerzita Hradec Králové, Pedagogická fakulta. [2011-09-19-2011-09-20].
- [2] VALA, R., LITSCHMANNOVÁ, M., VALOVÁ, M. *Prezentace: Zatížení extraligových hráčů basketbalu kategorie U-16 v utkání*. Hradec Králové: Univerzita Hradec Králové, Pedagogická fakulta. [2011-09-19-2011-09-20].